

On the robustness of the $H\beta$ Lick index as a cosmic clock in passive early-type galaxies

Alice Concas,^{1,2,3*} L. Pozzetti,³ M. Moresco^{2,3} and A. Cimatti²

¹*Excellence Cluster Universe, Boltzmannstr. 2 D-85748 Garching, Germany*

²*Dipartimento di Fisica e Astronomia, Università degli Studi di Bologna, V.le Berti Pichat, 6/2 - 40127, Bologna, Italy*

³*INAF - Osservatorio Astronomico di Bologna, via Ranzani 1, I-40127, Bologna, Italy*

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

We examine the $H\beta$ Lick index in a sample of ~ 24000 massive ($\log(M/M_\odot) > 10.75$) and passive early-type galaxies extracted from SDSS at $z < 0.3$, in order to assess the reliability of this index to constrain the epoch of formation and age evolution of these systems. We further investigate the possibility of exploiting this index as “cosmic chronometer”, i.e. to derive the Hubble parameter from its differential evolution with redshift, hence constraining cosmological models independently of other probes. We find that the $H\beta$ strength increases with redshift as expected in passive evolution models, and shows at each redshift weaker values in more massive galaxies. However, a detailed comparison of the observed index with the predictions of stellar population synthesis models highlights a significant tension, with the observed index being systematically lower than expected. By analyzing the stacked spectra, we find a weak $[\text{NII}]\lambda 6584$ emission line (not detectable in the single spectra) which anti-correlates with the mass, that can be interpreted as a hint of the presence of ionized gas. We estimated the correction of the $H\beta$ index by the residual emission component exploiting different approaches, but find it very uncertain and model-dependent. We conclude that, while the qualitative trends of the observed $H\beta$ - z relations are consistent with the expected passive and downsizing scenario, the possible presence of ionized gas even in the most massive and passive galaxies prevents to use this index for a quantitative estimate of the age evolution and for cosmological applications.

Key words: galaxies: general – galaxies: evolution – galaxies: formation – galaxies: stellar content – galaxies: fundamental parameters – cosmology: cosmological parameters –

* E-mail: alice.concas@tum.de (KTS)

1 INTRODUCTION

Early-type galaxies (ETGs) are perfect candidates to investigate both galaxy formation and evolution theories and, at the same time, to put constraints on the expansion history of the Universe. Many recent observations show that most massive galaxies contain the oldest stellar populations up to $z \sim 1-2$ (e.g. [Spinrad et al. 1997](#); [Cowie et al. 1999](#); [Heavens et al. 2004](#); [Thomas et al. 2005](#); [Cimatti et al. 2008](#); [Thomas et al. 2010](#); [Moresco et al. 2011](#)) and that only few percent of their current stellar mass was formed after $z \sim 1$. Moreover, studies on the stellar mass function at different redshifts (e.g. [Bundy et al. 2005](#); [Borch et al. 2006](#); [Bundy et al. 2006](#); [Ilbert et al. 2010](#); [Pozzetti et al. 2010](#); [Maraston et al. 2013](#)) showed that most massive galaxies ($M \sim 10^{11} M_{\odot}$) are characterized by an increase of their stellar mass density at $z \sim 1$ (e.g., [Ilbert et al. 2010](#) based on the COSMOS 2 deg² field) and a subsequent negligible evolution in number density from $z \sim 1$ to the present ([Cimatti et al. 2006](#); [Pozzetti et al. 2010](#); [Moresco et al. 2013](#)). All these observational evidences suggest that these massive, red and passively evolving galaxies represent the oldest objects in the Universe below $z \sim 1$. For this reason, a study based on the variation of ETGs properties, such as age, chemical composition, dynamic, at different redshifts, could rightly help to expand our current knowledge of the physical process that drive the formation and evolution of galaxies. Moreover, as firstly pointed out by [Jimenez & Loeb \(2002\)](#), the differential age evolution of ETGs can be also used to put several cosmological constraints. In particular, by estimating the relative differential age (Δt_{age}) between ETGs, formed at the same cosmic time, but observed at different redshifts (Δz), it is possible to determine the derivative of redshift with respect to cosmic time dz/dt .

This quantity is directly related to the cosmological expansion history described by the Hubble parameter: $H(z) = -\frac{1}{1+z} \frac{dz}{dt}$, which can be used to estimate many cosmological parameters such as H_0 , $\Omega_{0,M}$, $\Omega_{0,\Lambda}$ and w_{Λ} . This technique, known as “cosmic chronometers” method, has been used recently in several works and obtained promising results in the estimate of the Hubble parameter and of the dark energy equation of state ([Simon et al. 2005](#);

[Stern et al. 2010](#); [Moresco et al. 2012a,b](#); [Zhang et al. 2012](#); [Moresco et al. 2016b](#)). The basic problem of this method is the fact that it relies on the measure of stellar population age, which presents, when estimated from a fit of the spectrum and/or spectral energy distribution (SED), significant degenerations with other parameters, such as metallicity, star formation history (SFH) and dust content. An alternative method to mitigate these degeneracies is to identify a particular spectral feature that is sensitive to the aging of the stellar population. An example of this approach is provided by [Moresco et al. \(2011, 2012a\)](#), [Moresco \(2015\)](#) and [Moresco et al. \(2016a\)](#), in which the expansion history of the Universe is determined by studying the redshift dependence of the 4000 Å break. Despite the high potential of the method, the break at 4000 Å is particularly sensitive not only to the age, but also to the variation of the metallicity of the stellar population, which has to be taken into account as a systematic error in the estimate of cosmological parameters. It is evident the need to find a new spectral feature that minimize the dependence systematic and able to translate the amount dz/dt in the observable $dz/d(\text{feature})$ and then able to determine the cosmological parameter with higher accuracy.

In the effort of exploiting reliable age indicators, we present a new study of ETGs relative age evolution based on the variation of the H β Lick index with redshift. This index, firstly introduced by [Burstein et al. \(1984\)](#) and redefined by [Worthey et al. \(1994\)](#), has been identified as the one with the maximum dependence on the variation of the age of the stellar population, and the minimal dependence on metallicity and chemical composition ([Trager et al. 1998](#); [Lee & Worthey 2005](#); [Thomas et al. 2011](#)). In detail, we measure the H β Lick index in a sample of passive galaxies selected from the Sloan Digital Sky Survey (SDSS), to check if this particular spectral feature is actually able to provide a reliable differential dating of galaxy stellar ages, in order to derive constraints both on galaxy evolution and formation and on the expansion history of the Universe. This paper is organized as follows. In Section 2 we introduce the data sample. We describe the physical parameters of the passive galaxies sample, and our method for the H β

Lick measurement. In Section 3 we present the results about the $H\beta$ – z and $H\beta$ –mass relations and compare them with the theoretical predictions of the stellar population synthesis models. We discuss the effect of possible contamination by the emission lines in the observed $H\beta$ absorption line in the Section 4, and finally in Section 5 we list our conclusions. Throughout the paper we assume a $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_{0,M} = 0.25$ and $\Omega_\Lambda = 0.75$ cosmology.

2 DATA

The galaxies analyzed in this work are selected from the spectroscopic catalogue of the Sixth Data Release of Sloan Digital Sky Survey (SDSS-DR6). The SDSS-DR6 provides photometry in the u, g, r, i, z bands and the spectrum of about 800000 galaxies, extracted with Petrosian magnitude $r < 17.77$. Each spectrum is measured from 3800 to 9200 Å with a resolution $\lambda/\Delta\lambda \approx 1800 - 2000$ ¹. The SDSS data are provided in vacuum wavelengths. To be meaningfully compared with atomic transitions, we converted them in the air system². To have a wider photometric coverage, we decided to analyze the sample provided by Moresco et al. (2011), where SDSS galaxies are matched with the Two Micron All Sky Survey (2MASS), with photometry in the J, H and K bands. This wider coverage allows a better accuracy in the stellar mass estimates from the SED-fitting, and a more precise selection based on photometric data (see Sec. 2.1).

2.1 The sample of passive ETGs

This study strongly relies on the selection of a sample of passively evolving galaxies which, having assembled most of their mass at high redshifts, are able to trace homogeneously the age-evolution of the Universe. To select the purest possible sample of passive galaxies, with no evidence of recent

episodes of star formation, we followed the approach described in Moresco et al. (2011), and selected galaxies with:

- **no strong emission lines.** Starting from the equivalent width measurements provided by the analysis of the MPA-JHU DR7³, we select only the objects with a restframe equivalent width $\text{EW}(H\alpha) > -5 \text{ Å}$ and $\text{EW}([OII]\lambda 3727) > -5 \text{ Å}$, where, by convention, the measures are given in negative values when in emission.

- **spectral energy distribution matching the reddest passive ETG templates.** Following the approach of Zucca et al. (2006), the SED of each galaxy has been compared with a library of 62 empiric SEDs described in Ilbert et al. (2006), comprising galaxies with both old stellar populations and intense star formation activity. The galaxies are divided according to their SED, and we select those galaxies matching an early-type template, corresponding to the four reddest templates.

With this selection, we obtain a sample of ~ 100000 galaxies. For more details, we refer to Moresco et al. (2011).

In order to investigate the evolution of our sample, we study their physical properties, focusing in particular on the stellar mass, stellar metallicity and velocity dispersion. We use the stellar mass estimates from Moresco et al. (2011), obtained by a multi-colour SED-fitting procedure (8 photometric bands from u to K). We apply a further mass cut, selecting only the most massive galaxies with a stellar mass $\log(M/M_\odot) \geq 10.75$. This choice is made to select only the most massive objects that, according to the downsizing scenario, should correspond to the oldest objects (Thomas et al. 2010). In the *mass-downsizing* scenario, the most massive galaxies were formed earlier and over a shorter period than those in less massive galaxies. For this reason, we decide to split our sample in four narrow mass bins ($\Delta \log(M/M_\odot) = 0.25$), to have a more homogeneous sampling of the redshift of formation: $10.75 < \log(M/M_\odot) < 11$ (BinI), $11 < \log(M/M_\odot) < 11.25$ (BinII), $11.25 < \log(M/M_\odot) < 11.5$ (BinIII) and $\log(M/M_\odot) > 11.5$ (BinIV) (see Fig. 1). Furthermore, for this sample we have a measure of the Star Formation History (SFH) estimated by Moresco et al. (2011) with

¹ <http://www.sdss3.org/dr9/scope.php#specstats>

² The conversion relation from vacuum wavelengths to air wavelengths, proposed in Morton et al. (1991), is the following:

$$\lambda_{AIR} = \frac{\lambda_{VAC}}{\left(1.0 + 2.735182E-4 + \frac{131.4182}{\lambda_{VAC}^2} + \frac{2.76249E8}{\lambda_{VAC}^4}\right)} \quad (1)$$

with λ_{AIR} and λ_{VAC} expressed in Å.

³ <http://www.mpa-garching.mpg.de/SDSS/DR7/>

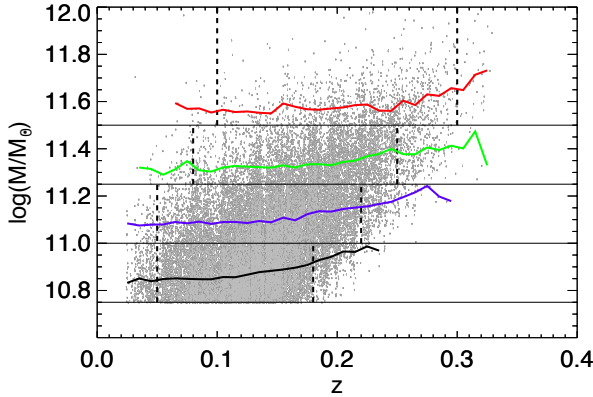


Figure 1. Stellar mass–redshift distributions of the four mass subsamples (red, green, blue and black curves respectively for the BinIV, BinIII, BinII and BinI, from top to bottom). The black dotted vertical lines indicate the redshift division of the sample (see Tab. 1 for the stellar mass median values). (A color version of this figure is available in the online journal.)

an exponential delayed SFH: $SFR(t, \tau) \propto (t/\tau^2) \exp(-t/\tau)$ (with τ in the range of values $0.05 \leq \tau \leq 1$ Gyrs). The median value of the our total sample is $\tau = 0.2$ Gyrs.

In Fig. 1 we present the mass-redshift distribution. It is evident that the median mass increases with the redshift due to magnitude limit effect of the survey. Therefore, to avoid possible biases, we select the galaxies having $z \leq 0.18$, $z \leq 0.22$, $z \leq 0.25$ and $z \leq 0.3$, respectively for the BinI, BinII, BinIII and BinIV. In this way, we limit the difference in the median mass along the redshift range to 0.1 dex on average (see Fig. 1).

As mentioned in the previous section, the $H\beta$ Lick index is the less sensitive to metallicity variations compared to all other indices of the Lick system. However, in the wavelength range where the index is defined, there may be weak metal absorption lines and, for this reason, the metallicity dependence of the $H\beta$ index is not totally negligible. The sample has been therefore cross-matched with the SDSS-DR4 subsample obtained by Gallazzi et al. (2005), for which the stellar metallicities of the galaxies are estimated from the simultaneous fit of several spectral features (D4000, $H\beta$, $H\delta_a + H\gamma_a$, $[Mg_2Fe]$ and $[MgFe]'$). We study the metallicity-redshift relation. In order to minimize the possible metallicity effect, we made a further cut in redshift, selecting only galaxies with redshift $z \geq 0.05$ for BinI and BinII, $z \geq 0.08$ for BinIII, and $z \geq 0.1$ for BinIV. The median metallicity of

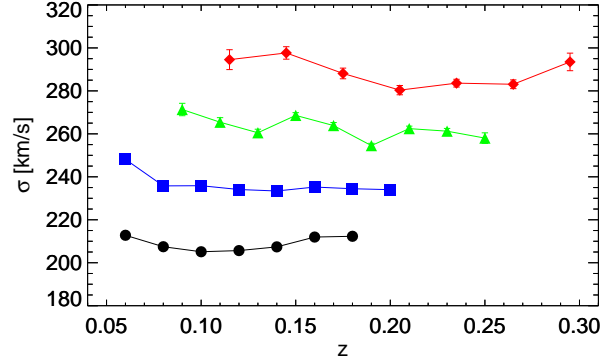


Figure 2. Velocity dispersion–redshift relations, averaged in bin of redshift, for the different mass subsamples, BinIV, BinIII, BinII and BinI from top to bottom (see Tab. 1 for the median values). (A color version of this figure is available in the online journal.)

the final sample is $Z/Z_\odot \sim 1.2$ (with a total spread of about 0.3 within the total redshift range analyzed and of ~ 0.2 in narrow redshift bins $\Delta z \sim 0.02 - 0.03$), showing only a slight variation with redshift, in agreement with the recent metallicity measurements obtained by total spectral fitting of our stacked spectra (Citro et al. 2015). We note, further, that there is no significant difference in the median metallicity of the different mass subsamples.

In this paper, we adopt the velocity dispersion derived by the Princeton group⁴. Figure 2 shows the velocity dispersion relations for all mass subsamples. We find that the velocity dispersion is constant in each subsample (see Tab 1 for the median values).

Finally, the observed optical spectrum can be contaminated by the presence of night sky emission lines (see Hanuschik 2003). Particularly the night sky emission lines at 5577 Å ([OI]) and at 5889.950, 5895.920 (NaI) Å are very strong, and leave significant contamination in $H\beta$ measurement.

In order to avoid this source of contamination, we decide to excluded the galaxies having a redshift between $0.14 < z < 0.155$ and $0.20 < z < 0.22$, which are the redshifts for which those lines fall into the $H\beta$ definition range.

After all the cuts adopted, the final sample consists of ~ 23914 galaxies. The median values and the respective er-

⁴ The data are available on the website <http://spectro.princeton.edu/>

	redshift z	mass range $\log(M/M_\odot)$	median mass $\log(M/M_\odot)$	median velocity dispersion σ [km s ⁻¹]	# galaxies
BinI	0.05 – 0.18	10.75 – 11.0	10.87 ± 0.001	207.3 ± 0.3	11631
BinII	0.05 – 0.22	11.0 – 11.25	11.10 ± 0.001	234.5 ± 0.4	7990
BinIII	0.08 – 0.25	11.25 – 11.5	11.34 ± 0.001	259.9 ± 0.6	3314
BinIV	0.1 – 0.3	> 11.5	11.58 ± 0.002	286.1 ± 1.1	979
Tot	0.05 – 0.3	> 10.75	11.01 ± 0.001	226.7 ± 0.2	23914

Table 1. Redshift range, mass range, median mass and median velocity dispersion of the passive galaxies in different mass subsamples.

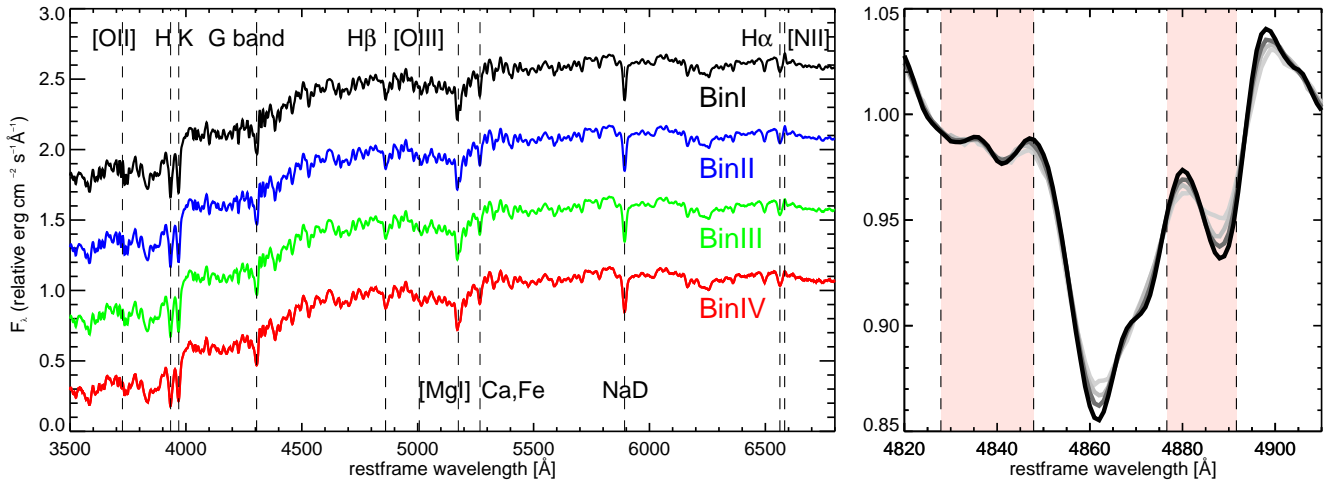


Figure 3. Median stacked spectra of the four passive galaxies mass subsamples. (from top to bottom respectively for the BinI, BinII, BinIII and BinIV, left panel). In the range between 3500 – 6800 Å there are not significant emission lines ([OII], [OIII], H α and [NII] λ 6584). The H β line show a typical absorption feature (right panel) in all the mass bins (from black to light gray for BinI, BinII, BinIII and BinIV respectively). The spectra are normalized between 5000 and 5500 Å restframe. (A color version of this figure is available in the online journal.)

rors⁵; of mass and velocity dispersion for the different mass subsamples are reported in Table 1. In the same table, we also report the number of galaxies in each bin.

Fig. 3 shows the median stacked spectra for the four mass subsamples. From the left panel, it is evident that all spectra show typical features of a passive population, with (i) absence of emission lines ([OII], H β , [OIII], and H α), (ii) strong absorption features (H β , H α , G band, Ca H and K), and (iii) strong D4000 break. We notice that there is still present a weak emission at 6584 Å [NII], and that the amplitude of this feature is mass-dependent, being most evident in the less massive subsample (BinI). For more details,

see Section 4. We highlight the region near H β line (right panel); all the four spectra show a typical and unperturbed absorption profile.

2.2 H β index

We measure the H β Lick absorption line index for all 23914 galaxies. We implement an IDL (Interactive Data Language⁶) code in order to perform measures coherently on observational data and models. We use the Lick H β index definition by Worthey et al. (1994); the bandpasses are shown in Fig. 3. In the standard Lick system, the indices are defined on low resolution spectra (~ 9 Å FWHM in H β region). However, the modern spectroscopic surveys, and in particu-

⁵ The error on the median are evaluated as the *median absolute deviation* $/\sqrt{N}$, $MAD = 1,482 * median(|x - median(x)|)$ (see Hoaglin et al. 1983).

⁶ <http://www.exelisvis.com/ProductsServices/IDL.aspx>

lar the SDSS, provide spectra with a much higher resolution ($\sim 2.4 - 2.7 \text{ \AA}$ FWHM), comparable with the new stellar populations synthesis models (for example [Bruzual & Charlot 2003](#) and [Maraston & Strömbäck 2011](#)). In order to measure coherently this feature in the models and observations, we decide to perform the analysis directly on the observed spectra, without altering the resolution.

Before applying our code to real data, we test its reliability to reproduce the $H\beta$ Lick values tabulated in the BC03 templates. We find a very good agreement between the two measurements with a negligible median offset of 2 per cent. We also compare our measurements with those derived by the MPA-JHU group for the same galaxy sample (<http://www.mpa-garching.mpg.de/SDSS/DR7/>). We find a good agreement between the values, with a mean difference of -0.03 \AA and a standard deviation of 0.033 \AA , which is comparable with the average measured error. Therefore, we find that the two independent methods are consistent, and without any significant biases.

3 TRENDS OF $H\beta$

We derive the median $H\beta$ -redshift relations in each mass subsample; we use narrow redshift bins ($\Delta z \sim 0.02 - 0.03$) for the three most populated mass subsamples, while in the more massive subsample we use $\Delta z \sim 0.04$ because of the lower statistic with respect to the other samples.

The results are shown in Fig. 4. From this figure, we note two main trends:

- a clear $H\beta$ -redshift relation for each subsample, with an increase in the index strength of $\sim 10\%$ along the redshift range.
- a clear trend with mass; i.e., at each redshift more massive galaxies present a lower $H\beta$ index with respect to less massive ones ($\sim 9\%$).

It is interesting to note that both effects can easily be seen even by direct inspection of the median spectra. As an example, we show the stacked spectra at different redshifts for the sample with $11 < \log(M/M_\odot) < 11.25$ (BinII, see Fig. 5, left panel), from which it is evident a deeper absorption line with redshift. The trend with the mass is also

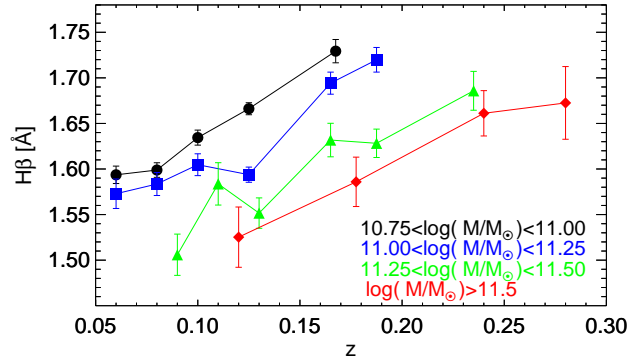


Figure 4. Median $H\beta$ -redshift relation for passive galaxies in different mass subsamples (circles, squares, triangles and diamonds for BinI, BinII, BinIII and BinIV respectively). (A color version of this figure is available in the online journal.)

visible in the stacked spectra of galaxies at fixed redshift, with most massive galaxies showing a smaller absorption line than lower-mass ones (see Fig. 5, right panel).

Both these trends are qualitatively consistent with a general aging of the stellar population and a mass-age relation (“mass-downsizing”), also found by [Thomas et al. \(2010\)](#). Our results also concur with a recent work of [Choi et al. \(2014\)](#). By modeling the full optical spectrum of a quiescent galaxy sample at low and intermediate redshift ($0.07 < z < 0.09$ from the SDSS and $0.1 < z < 0.7$ from AGN and Galaxy Evolution Survey, AGES), [Choi et al. \(2014\)](#) found that the best-fit SSP-equivalent age is higher for the most massive galaxies and it increases with decreasing redshift at fixed stellar mass. This correspondence between two different methods highlights the potential of our $H\beta$ line as a tracer of aging in the stellar population. We test this interpretation in Sec. 3.1.

3.1 Sensitivity to metallicity and velocity dispersion

The $H\beta$ index is not totally immune to metallicity (Z/Z_\odot) effects for the integrated stellar population ([Worthey et al. 1994](#)). Furthermore, this index is sensitive to the galaxy velocity dispersion (σ). To exclude the possibility that the trends found are due to a σ or Z/Z_\odot variation, we study the impact of these parameters using stellar population synthesis (SPS) models. We consider [Bruzual & Charlot \(2003\)](#) and [Maraston & Strömbäck \(2011\)](#) models (hereafter BC03

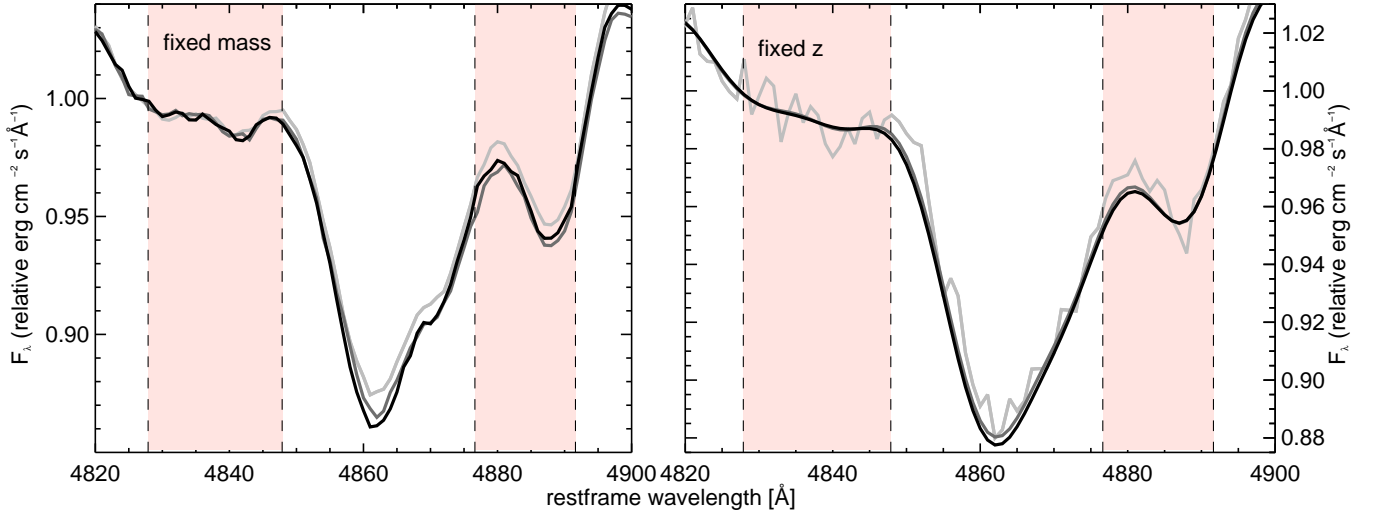


Figure 5. Median stacked spectra in the $H\beta$ line region. Left panel: stacked spectra with the same mass ($11 < \log(M/M_\odot) < 11.25$, BinII) and different mean redshift $z = 0.08, 0.10$ and 0.1875 (from light gray to black, respectively). Right panel: stacked spectra with the same redshift, $z = 0.125$, for different mass subsamples (black, gray and light gray for BinI, BinII, and BinIV respectively). For illustrative purposes the BinI and BinII spectra are smoothed to the common velocity dispersion of BinIV.

and MaStro, respectively). BC03 models are available at a resolution of 3 \AA FWHM in the wavelength range between $3200 - 9500 \text{ \AA}$; MaStro models are available with a resolution of $\sim 2.3 \text{ \AA}$ FWHM⁷ between $3525 - 7500 \text{ \AA}$. Both models share a resolution very similar to the one of SDSS spectra ($\approx 1800 - 2000$ between 3800 to 9200 \AA). This is particularly useful since it allows a direct comparison between the observed spectra with the theoretical SED without having to modify the spectral resolution.

These models are used to create libraries of synthetic spectra with different velocity dispersions ($\sigma = 0, 200, 250$ and 300 km s^{-1}) and different metallicity ($Z/Z_\odot = 1, 1.4, 1.5$ obtained by interpolating quadratically between $Z/Z_\odot = 0.4, 1, 2.5$ available in BC03 and $Z/Z_\odot = 0.5, 1, 2$ in MaStro). The grid has been created with age $0.5 \leq t \leq 15 \text{ Gyr}$, with de-

layed exponential SFH, $\text{SFR}(t, \tau) \propto (t/\tau^2) \exp(-t/\tau)$, where τ is chosen in the range $0.05 \leq \tau \leq 1 \text{ Gyr}$.

3.1.1 Effects of the velocity dispersion

To check the effect of the velocity dispersion in our $H\beta$ Lick measures, we perform several tests. By using the theoretical spectra described in the previous section, we verify that an increase of velocity dispersion causes an apparent decrease of the observed $H\beta$ index strength only due to a broadening effect. We further estimate the maximum percentage difference in $H\beta$ between our mass bins which can be attributed to σ from the analysis of SPS models; we find that the percentage difference in $H\beta$ between a single stellar population (SSP) with $\sigma = 200 \text{ km s}^{-1}$ and $\sigma = 250 \text{ km s}^{-1}$ is $< 1.5\%$, and is $< 3\%$ between $\sigma = 200 \text{ km s}^{-1}$ and $\sigma = 300 \text{ km s}^{-1}$, for both models. We underline that given the range of σ of our data, these two estimates represent the maximum difference in $H\beta$ which can be attributed to a σ effect between our

⁷ More precisely, in Beifiori et al. (2011) it is shown that the spectral resolution of the MaStro models, based on the library stellar MILES is about 2.54 \AA FWHM, instead of the nominal 2.3 \AA

mass subsamples, and they are always lower than the mean percentage difference in our subsamples with similar σ differences. Indeed, the mean percentage difference in $H\beta$ median values between BinI and BinIII is $\sim 6\%$, and $\sim 9\%$ between BinI and BinIV (similarly, we get a difference of $\sim 4\%$ between BinII and BinIII, and $\sim 6\%$ between BinII and BinIV). Moreover, we note that all SPS models analyzed show that SEDs with different velocity dispersions have a difference in flux in the red pseudo-continuum region ($4876.625 - 4891.625$), but the stacked spectra of the same mass subsample do not show this trend (see Fig. 5).

For this reason, we exclude that the velocity dispersion by itself can cause the observed mass segregation between the $H\beta$ -z relation in different mass subsamples.

Finally, we verified that in our data the median velocity dispersion (see Table 1 for median values of BinI, BinII, BinIII and BinIV) shows almost no redshift evolution in each mass bin (see Fig. 2). For this reason, we exclude that the observed $H\beta$ index increase with redshift can be due to a velocity dispersion effect.

3.1.2 Effects of the metallicity

The models are also used to establish the effect in the $H\beta$ index due to the stellar metallicity. While there is no significant difference in the median metallicity of the different mass subsamples, each mass bin instead shows a slight decrease of the median metallicity with redshift.

We create a library of $H\beta$ indices as a function of age using stellar population models with different metallicity: 1, 1.4 and 1.5 Z/Z_\odot . The study of this $H\beta$ synthetic indices shows that even an extreme variation of metallicity, between $Z/Z_\odot = 1$ to $Z/Z_\odot = 1.4, 1.5$, implies a $H\beta$ variation of only $\sim 4-5\%$ on average, always lower than $H\beta$ trend observed. Moreover, to further check the metallicity effect, not only in the models but also on our data, we create a control sample of galaxies with a narrow metallicity range (e.g., $\Delta(Z/Z_\odot) = 0.5$ for the BinI). Even in this subsample we find a steady trend of growth in the $H\beta - z$ relations.

From these analyses, we find that the increase of the $H\beta$ equivalent width with redshift (at fixed mass) and decrease with mass (at fixed redshift) of passive galaxies cannot be

only due to σ or Z/Z_\odot variation, and hence can be mostly explained by an age-evolution effect, in particular with the age of the stellar population decreasing with increasing redshift, and increasing with increasing mass.

3.2 $H\beta$ -age: Comparison with SPS models

To interpret the evolution of $H\beta$ with redshift and its dependence on mass in the context of passive galaxies evolution, we proceed with a direct comparison between the observed data and the index values calculated on the SED provided by SPS models, hereafter theoretical $H\beta$. Both the observed and theoretical measurements are obtained by the same method, described in Section 2.2. The theoretical $H\beta$ curves as a function of age are transformed into $H\beta$ -z relations assuming a Λ CDM cosmology (see Section 1 for the parameters used), and a formation redshift. For our studies, we chose to probe different redshifts of formation for our galaxies, namely $z_F = 1, 2, 3, 4$ and ∞ . For this comparison we use two different stellar populations synthesis models, BC03 and MaStro and the theoretical models of $H\beta$ Lick indexes of Thomas et al. (2011) (hereafter TMJ), calculated by theoretically manipulating the index response functions. For BC03 and MaStro models, we adopt a Chabrier IMF instead the TMJ models have a Salpeter IMF. However, using BC03 models, we verify that a IMF variation does not affect significantly the $H\beta$ value. In particular, in the case of solar metallicity, the choice of a Salpeter IMF instead of a Chabrier causes an index decrease of $< 1\%$. Furthermore, in the BC03 and MaStro models we use a delayed exponentially SFH, in particular we use models with $\tau = 0.2$ Gyrs which correspond to the τ median value estimate for our sample (see Section 2.1). We investigate also the possible variation of $H\beta$ value with different τ values. We compare SPS models with $\tau = 0.05$ and $\tau = 0.2$ Gyrs, and we find that for SPS older than 6 Gyr the $H\beta$ Lick index decreases of $\leq 1\%$ in the SPS with $\tau = 0.05$ Gyrs to respect at SPS with $\tau = 0.2$ Gyrs. Since the median metallicity of the sample is $Z/Z_\odot \sim 1.2$, for the comparison with the data we interpolated quadratically the $H\beta$ values of all models in the three metallicities provided, in order to have an estimate of the index at $Z/Z_\odot = 1.2$. The BC03 and MaStro models are convolved

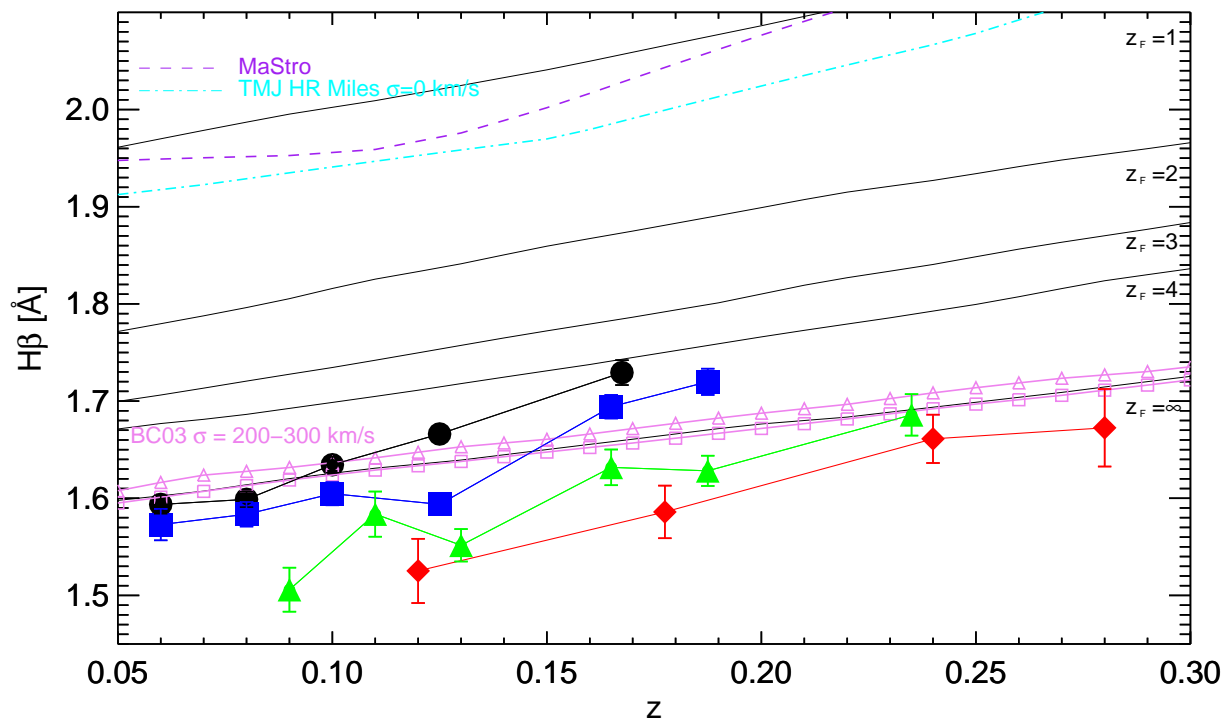


Figure 6. Comparison between the observed $H\beta$ - z relations (circles, squares, triangles and diamonds for BinI, BinII, BinIII and BinIV respectively) and different models. The black solid lines are the BC03 models with $\sigma = 230 \text{ km s}^{-1}$ $Z/Z_{\odot} = 1.2$ and different formation redshifts ($z_F = 1, 2, 3, 4$ and ∞). For $z_F = \infty$ we show also the models with $\sigma = 200$ and 300 km s^{-1} , open triangles and open squares respectively. The MaStro and TMJ models are shown for $z_F = 1$ (dashed and dash and dot lines, respectively). (A color version of this figure is available in the online journal.)

with a velocity dispersion $\sigma = 230 \text{ km s}^{-1}$, comparable to the median value of our sample. In Fig. 6, we show as an example the observed $H\beta$ - z relations and the theoretical curves (BC03 models) for different formation redshifts. For MaStro and TMJ models we find similar results, as shown in the case of formation redshift $z_F = 1$.

From the comparison with theoretical models, as can be clearly seen from Fig. 6, we find that they do not reproduce the observed $H\beta$ median relations, since, the models predict higher $H\beta$ values than the median $H\beta$ measured on our samples, so that in most cases the observed values would require an age greater than the age of the Universe. For the most massive sample the minimum shift between data and models is of the order of $\Delta H\beta = 0.1$.

Since the models are built with the same metallicity and velocity dispersion found in the data, we exclude that this inconsistency is due to one of these two parameters. We verify that also by imposing an extreme velocity dispersion for our

sample ($\sigma = 300 \text{ km s}^{-1}$) the models do not reproduce the observations. Regarding the metallicity, we instead find that the $H\beta$ - z relations can be reproduced only by considering models with metallicity approximately $Z/Z_{\odot} = 2$, that are inconsistent with the metallicity estimated for this sample.

Furthermore, comparing the metallicity estimated by Thomas et al. (2010) for a morphological selected sample of ETGs at $0.05 \leq z \leq 0.06$, we notice that for a subsample with velocity dispersion similar to our values they estimate a slightly higher metallicity than those considered in this work (Gallazzi et al. 2005). In particular, for $\sigma = 207 \text{ km s}^{-1}$ (median value of the less massive sample, BinI) Thomas et al. (2010) find a metallicity $Z/Z_{\odot} \sim 1.5$, and for the median value of the total sample, $\sigma = 230 \text{ km s}^{-1}$, they estimate a metallicity of the order of $Z/Z_{\odot} \sim 1.6$ (for more details, refer to Thomas et al. 2010). These higher metallicity values are also confirmed in a more recent work by Citro et al. (2015). In this study the authors find a median $Z/Z_{\odot} \sim 1.450 \pm 0.075$.

Even these higher values are not enough to completely reconcile models with data.

4 INVESTIGATING THE ORIGIN OF THE OFFSET

In this section, we further investigate possible causes of the offset between the model predictions and the measures of the $H\beta$ Lick index showed in Section 3.2. In particular, we explore the possibility that our sample, despite the selection criteria (see Section 2.1), is still contaminated by an emission-line component. This potential emission component would contaminate the measurement of the $H\beta$ index by filling the line and producing a less intense absorption feature, consistent with the recent results showed by [Serven & Worthey \(2010\)](#) for a SDSS low redshift ($0.06 < z < 0.08$) galaxy sample.

4.1 Improving sample selection: the [NII] emission-line contamination

As discussed in Section 2, we select only the galaxies that not showing significant emission lines in $H\alpha$ and [OII] ($EW < 5 \text{ \AA}$). We check that also cutting the sample by adopting a more stringent $H\alpha$ limit ($S/N < 3$), the offset between the $H\beta$ Lick index in our data and models is not removed. However, a more detailed analysis of the stacked spectra shows the presence of a weak [NII] emission at 6584 \AA (see Fig. 3). Since the first ionization potential of the nitrogen is very similar to that of hydrogen, the presence of an emission line in [NII] $\lambda 6584$ may indicate the presence of an emission component also in the H lines. In order to study the possible emission line contamination in our sample, we use the $EW([NII]\lambda 6584)$ values provided by the MPA-JHU group. From this analysis, about 43% of the total sample have an equivalent width $EW([NII]\lambda 6584)$ with $S/N > 3$, consistent with the clearly visible [NII] $\lambda 6584$ emission line in the median spectra (see Fig. 3 left panel). The median and dispersion⁸ equivalent width of the total sample

are $EW([NII]\lambda 6584) = -0.56 \pm 0.58 \text{ \AA}$, in agreement with the equivalent width measured on the median stacked spectrum ($EW([NII]\lambda 6584) = -0.68 \text{ \AA}$).

Moreover, by analyzing separately each mass subsample, we identify a relationship of the [NII] $\lambda 6584$ emission line as a function of stellar mass. In Fig. 7, bottom right panel, we show the stacked spectra of the four mass subsamples, from which it is evident that the [NII] $\lambda 6584$ emission line increases with decreasing of the median mass. This evidence is confirmed also by analyzing the $EW([NII]\lambda 6584)$ - z relations: we note that the $EW([NII]\lambda 6584)$ median values are almost constant with redshift, and they decrease with increasing mass (see Fig. 8).

To verify if the presence of a [NII] $\lambda 6584$ emission is related to the presence of $H\beta$ emission line, we analyze the galaxies most affected by this contamination. We select the galaxies with a [NII] $\lambda 6584$ clearly detected, i.e. with $S/N([NII]\lambda 6584) > 10$; in this way, we select 1145 galaxies⁹ (5% of the total sample). Figure 7 shows the median spectrum relative to this subsample. It can be seen that, differently from the stacked spectrum of the global sample (see Fig. 3), in this spectrum, together with a strong [NII] $\lambda 6584$, there are clearly detectable several emission lines, i.e. [OII], [OIII] $\lambda 5007$, [NII] $\lambda 6548$, $H\alpha$, [SII] $\lambda 6717$ and [SII] $\lambda 6731$ (Fig. 7). Nevertheless, also in this case, in the $H\beta$ region, there is not a clear emission component (Fig. 7 bottom left panel). The $H\beta$ absorption line only shows a particular shape, different from that expected for a simple absorption, probably caused by an overlapping with a weak emission line contamination.

Following this result, we decide to apply a further cut to our sample, by selecting only the galaxies with a [NII] $\lambda 6584$ line detected with $S/N([NII]\lambda 6584) < 3$; this sample is composed by 13626 galaxies ($\sim 57\%$ of the total sample). We find that, even for this sample there is a residual [NII] $\lambda 6584$ line, with a median and dispersion (MAD) $EW([NII]\lambda 6584) \sim -0.26 \pm 0.27 \text{ \AA}$. The residual emission can

⁸ The dispersion is evaluated with the *median absolute deviation*, *MAD*.

⁹ In this work, we decide to cut the sample above this limit ($S/N([NII]\lambda 6584) > 10$) because enabled us to select only those objects in which the [NII] line is detected effectively while maintaining a good statistic (1145 galaxies).

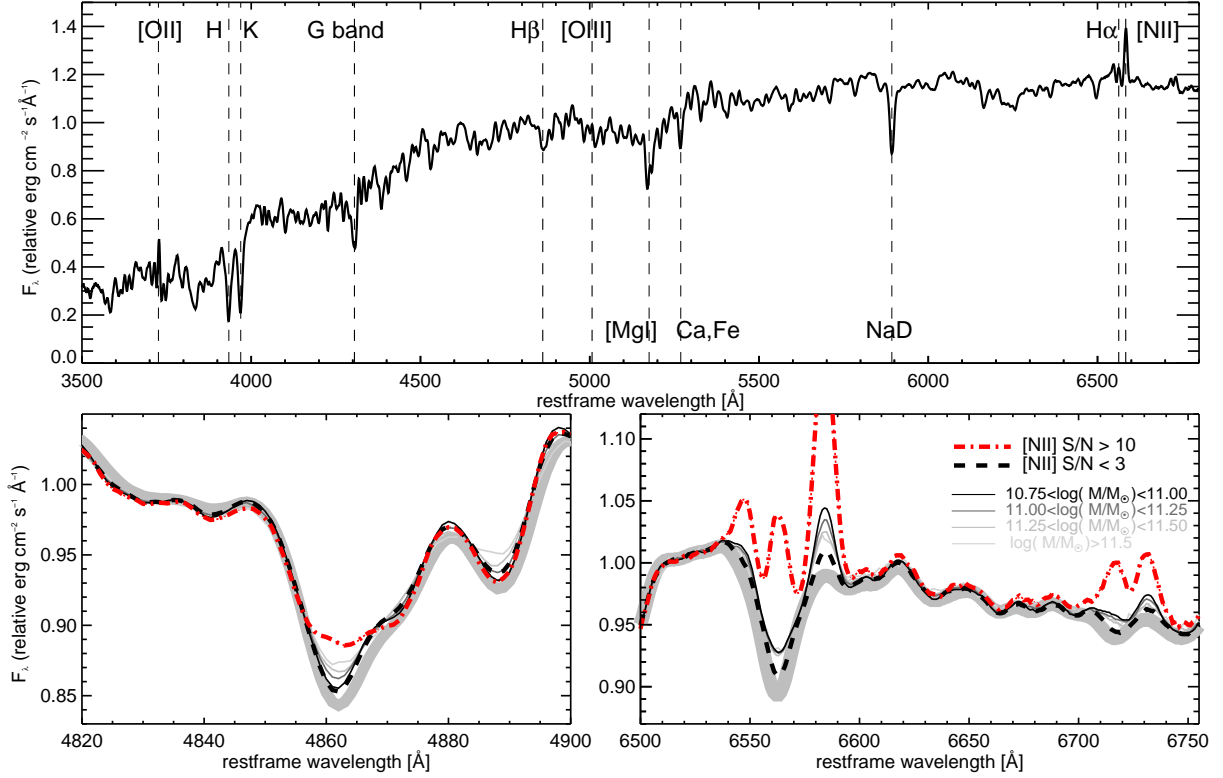


Figure 7. Top panel: stacked spectrum of the sample with $[NII]\lambda 6584$ S/N > 10 (1145 galaxies), it shows some evident emission lines: [OII], [OIII], H α and $[NII]\lambda 6584$. Bottom panels: comparison between the $[NII]\lambda 6584$ S/N > 10 stacked spectrum (red dash-dot lines), $[NII]\lambda 6584$ S/N < 3 stacked spectra (black dashed lines) and the stacked spectra of the original sample in different mass subsamples (from black to light gray for BinI, BinII, BinIII and BinIV respectively). In the H β region (left panel) are not evident emission lines, instead, in the H α region there are evident emissions ($[NII]\lambda 6584$, H α , $[NII]\lambda 6584$, [SII] $\lambda 6717$ and [SII] $\lambda 6731$). The gray area is the SSP solar model with age = 10 Gyrs, $\sigma = 250$ km s⁻¹ calculated by BC03 model. (A color version of this figure is available in the online journal.)

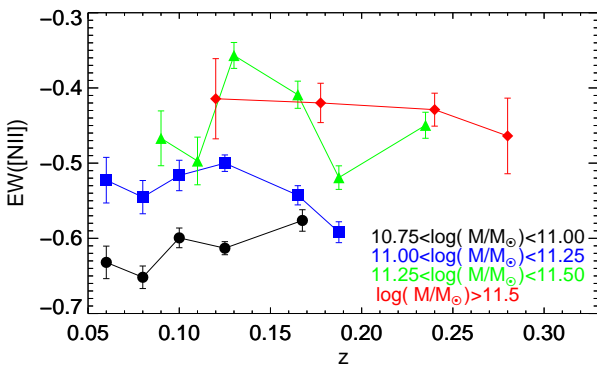


Figure 8. $[NII]\lambda 6584$ -redshift relations for the four mass subsamples (circles, squares, triangles and diamonds for BinI, BinII, BinIII and BinIV respectively). (A color version of this figure is available in the online journal.)

be seen from the comparison between the stacked spectrum of this sample and a pure absorption SPS model (see Fig. 7 bottom panels).

On this new subsample with $S/N([NII]\lambda 6584) < 3$, we

calculate also the median $H\beta$ - z relations at different mass, the results are shown in Fig. 9. We note that the cut in $[NII]\lambda 6584$ produces a general increase in median values of the median $H\beta$ index (on average to ~ 0.05 Å for the BinI and BinII, ~ 0.02 for the BinIII and ~ 0.01 for the BinIV). The offset found between data and observations of the total sample is reduced. However, there is still a significant disagreement between data and theoretical $H\beta$ - z curves, especially for the more massive sample (Bin IV). However, despite this discrepancy, we find similar $H\beta$ - z and $H\beta$ -Mass relationships as seen in the total sample, but with slightly different slopes and normalizations. These variations make impossible to use the median $H\beta$ values for an accurate estimate of the absolute and differential age evolution of the stellar populations with the redshift. Consequently, it is not possible to place constraints on the redshift of formation of

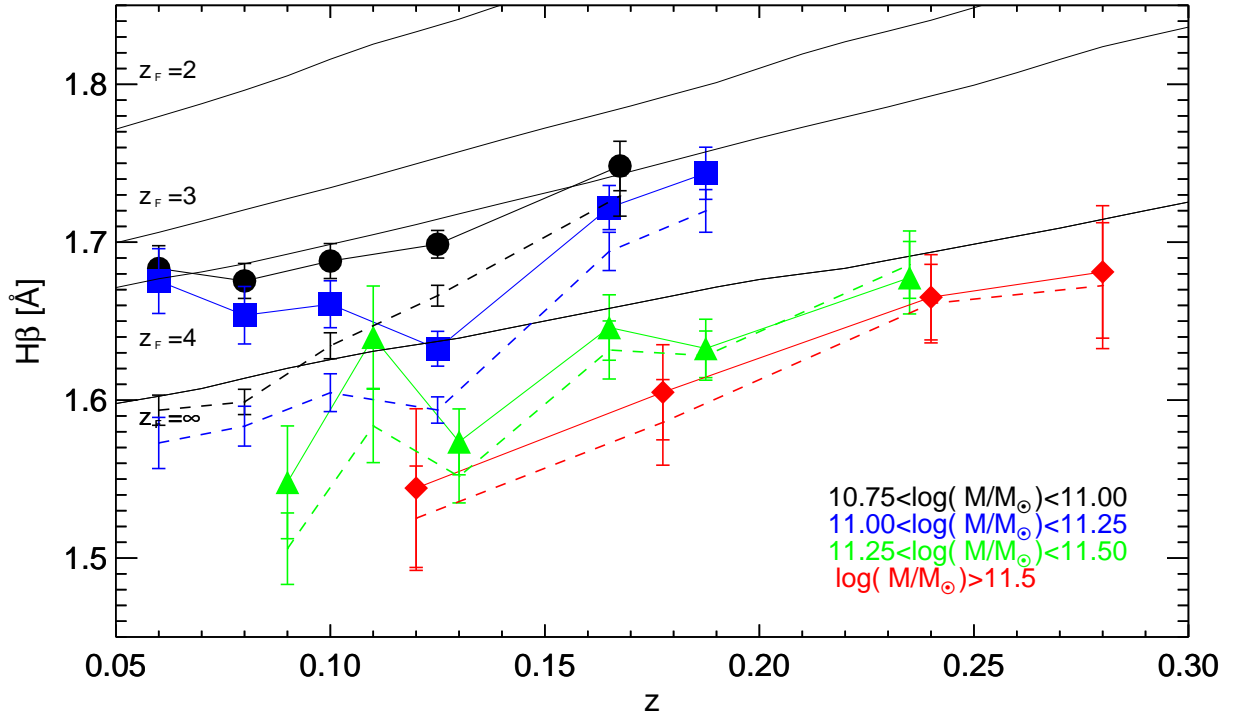


Figure 9. Comparison between $H\beta$ - z relations (circles, squares, triangles and diamonds for BinI, BinII, BinIII and BinIV respectively) of the total sample (dotted curves) and the sample with $EW([NII]\lambda 6584) S/N < 3$ (continuous curves). (A color version of this figure is available in the online journal.)

the galaxies and even the use of the $H\beta$ - z relation in the method of the cosmic chronometers appears inadequate.

4.2 Methods to remove nebular emission

The SDSS galaxies spectra are extensively used for many purposes, such as stellar mass and star formation history estimates (Kauffmann et al. 2003), age and metallicity estimates in the local universe (Gallazzi et al. 2005), environment effects in the galaxy formation (Thomas et al. 2010), and many other. In these previous works, before interpreting the observed spectra, and therefore before using the stellar absorption-line indices, standard procedures are followed to remove the contamination by nebular emission lines. In this section we try to apply these cleaning methods both in the individual and in the stacked galaxies spectra.

The method proposed by Tremonti et al. (2004), for example, makes a non-negative least-squares fit of the emission line free-observed regions of the spectrum, using a spectral library built using BC03 models. Then, the fitted spec-

trum is subtracted from the observed spectrum, and the residuals can be fitted to Gaussian broadened emission lines templates. Finally, the fitted emission lines are subtracted from the original observed spectrum to produce a “pure” absorption line spectrum (for further details, we refer to Tremonti et al. 2004). To verify if this method of decontamination is actually able to producing a clean sample in the $H\beta$ region, we use $H\beta$ values measured by the MPA-JHU DR7 group, which are available both for the original observed spectra that for the “pure” absorption line spectra. We select a sample consisting on galaxies without emission lines, i.e. with no correction for emission lines based on the method of Tremonti et al. (2004), which contains 19904 objects. We measure the median $H\beta$ - z relations also for this subsample, with the same procedure described in Section 3. The $H\beta$ - z relations obtained are consistent with those found for the subsample with $S/N([NII]\lambda 6584) < 3$. This suggests that, using objects with no emission lines detected by standard correction method, the offset between the observed $H\beta$

values and the predicted ones is not totally removed.

In addition, we test the presence of weak $H\beta$ emission line directly in the stacked spectra. We split our sample in terms of stellar mass and redshift as we done in Section 3 for the $H\beta$ values, and then we built the median spectra for each mass and redshift bin. In order to separate the stellar continuum from the nebular emission component, we decide to using a combination of the publicly available IDL codes: penalized pixel-fitting (pPXF), developed by (Cappellari & Emsellem 2004) and gas and absorption line fitting (GANDALF) write by (Sarzi et al. 2006). We measure the line-of-sight velocity distribution (LOSVD) by using pPXF code. Then, we perform the GANDALF analysis to convolve a set of input synthetic spectra with the previously kinematic and to fit the observed stacked spectrum simultaneously with the models and a Gaussian emission lines template. The result is a superposition of an optimal combination of the SSP templates with a set of Gaussians that represent the emission lines. Through the subtraction of the emission line spectrum from the observed one, we get the clean absorption line spectrum free from emission line contamination. First, we analyze the obtained $H\beta$ emission lines and then we explore the impact of this cleaning method in the observed spectra. In order to test the robustness of the emission lines extraction, we repeat the methodology by using three different spectral library: two based on the BC03 SSP ¹⁰ with different age $0.01 \leq t \leq 14$ Gyr (hereafter BC03_14Gyr) and a more extended age $0.01 \leq t \leq 20$ Gyr (hereafter BC03_20Gyr); and the third library built with MaStro models with and age $0.01 \leq t \leq 14$ Gyr. We find that, the recovered $H\beta$ emission lines are model dependent: the equivalent width is systematically higher for the BC03_20Gyr library than the BC03_14Gyr ($\Delta EW \sim 0.02 \text{ \AA}$) and the differences increase if we compare BC03 with MaStro models, $\Delta EW \sim 0.2 \text{ \AA}$. The median equivalent width are : $EW \sim -0.09 \pm 0.005$, $\sim -0.07 \pm 0.01$ and $\sim -0.27 \pm 0.01 \text{ \AA}$ for BC03_14Gyr, BC03_20Gyr and MaStro library, respectively. This result is consistent with the study showed

in Singh et al. (2013); their figure 4 shows that at low $H\beta$ fluxes there are differences in the flux extraction by using GANDALF code with SSP instead of the stellar templates. Therefore, the correction for very low emission in the absorption $H\beta$ line depends on the templates library used in the continuum fit. Furthermore, we find that, in all the three libraries, the $H\beta$ emission lines are detected with a very low confidence level, the amplitude over noise is always $A/N \leq 3$ (the noise is defined as the dispersion of fluctuations in the fit residuals).

Finally, we test the more recent emission correction for the hydrogen features proposed by Serven & Worthey (2010). Serven & Worthey (2010) derived emission corrections of the Balmer series Lick indices for a SDSS quiescent galaxy spectral sample (Graves et al. 2007) by comparing the $H\alpha$ -Mg b diagram from the SDSS stacked spectra with the measurements obtained for 13 Virgo galaxies. By using the same prescriptions showed in Serven & Worthey (2010), we re-calibrate the $H\alpha$ -Mg b diagram and then the $H\beta$ emission correction in our stacked spectra. We find that for the more massive sample (BinIV), the mean Mg b value is $4.23 \pm 0.15 \text{ \AA}$ and the mean $H\beta$ emission correction factor is on the order of $-0.17 \pm 0.08 \text{ \AA}$, that it is in good agreement with the emission values obtained with the GANDALF method mentioned before. However, we stress that, also in this case, the method is model-dependent, since the continuum correction was determined by using the Worthey et al. (1994) and Trager et al. (1998) models (see Section 2 of Serven & Worthey 2010). Moreover, rescaling the $H\alpha$ -Mg b relation from the SDSS values to the Virgo data could be age-dependent if the two samples have different mean ages. We conclude that for our sample, also in the stacked spectra, is very difficult measure a well detected and model-age-independent $H\beta$ emission line and than correct our observed spectra in order to obtain a "pure" absorption $H\beta$ line.

4.3 $H\beta$ self-diluted emission line

In order to have a qualitative estimate of the hidden $EW(H\beta)$ emission related to the observed $[NII]\lambda 6584$ emission line, we study the emission line ratios of our spectra in the diagnostic diagrams. We adopt the classical

¹⁰ The BC03 and MaStro stellar population templates have a Chabrier IMF and metallicity $Z/Z_{\odot} = 0.2, 0.4, 1, 2.5$ and $Z/Z_{\odot} = 0.5, 1, 2$, respectively.

(Baldwin et al. 1981) diagram, hereafter BPT diagram. The equivalent width measurements for [OIII], H β , H α and [NII] λ 6584 are taken by the MPA-JHU group analysis. In Fig. 10 we show the BPT diagrams for the subsample with all EWs detected with S/N>3, the subsample with EW([NII] λ 6584) with S/N>10, and the subsample with EW([NII] λ 6584) with S/N<3. In the first two cases, we find that the majority of the objects are located in the LINERs region (Low Ionization Nuclear Emission line Regions, Heckman 1980). For the first and the second sample, the median and dispersion for $\log(EW([NII])/EW(H\alpha))$ are -0.014 ± 0.159 and 0.050 ± 0.177 , and for $\log(EW([OIII])/EW(H\beta))$ are 0.058 ± 0.139 and 0.162 ± 0.312 , respectively. In the last subsample, the dispersion in the data is larger because of the lower signal to noise ratio (see Fig. 10). In this last case, the median value of the sample is located in the composite region ($\log(EW([NII])/EW(H\alpha)) \sim -0.150 \pm 0.341$ and $\log(EW([OIII])/EW(H\beta)) \sim 0.103 \pm 0.413$).

By using the median $EW([NII]\lambda 6584)/EW(H\alpha)$ values and the median $EW([NII]\lambda 6584)$ we estimate qualitatively the $EW(H\alpha)$ contribution. Then, we estimate the $EW(H\beta)$ by assuming no absorption by dust, and hence an emission line ratio, $H\alpha/H\beta \sim 2.86$ (Osterbrock et al. 1989), for electron density of $n = 100 \text{ cm}^{-3}$ and electron temperature $T_e = 10^4 \text{ K}$.

In the sample with $EW([NII]\lambda 6584)$ with S/N<3 the $EW([NII]\lambda 6584)$ is $\simeq -0.26 \pm 0.27 \text{ \AA}$. For this sample we expect an $EW(H\beta)$ emission in the range -0.28 to -0.06 \AA (by considering the $EW([NII]\lambda 6584)/EW(H\alpha)$ dispersion), or weaker in case of dust extinction.

The same conclusions can be drawn for the sample with the larger [NII] λ 6584 emission lines ($S/N([NII]\lambda 6584) > 10$). In this case, the median $EW([NII]\lambda 6584)$ is $-1.99 \pm 0.58 \text{ \AA}$, and the interval of variability of $\log(EW([NII])/EW(H\alpha))$ is smaller, $\sim 0.050 \pm 0.177$. Then, we expect an H β emission line equivalent width between -0.9 and -0.4 \AA in case of no absorption by dust, or weaker otherwise.

In order to verify whether the H β emission line values, expected in the two different samples, $S/N([NII]\lambda 6584) < 3$ and > 10 , can be effectively detected in the observed spec-

tra, we perform some simulations, described in the following section.

4.4 H β emission line simulations

Starting from the theoretical SED (in which only the absorption component is taken into account) of SPS models (BC03), we simulate the presence of an emission line by adding a Gaussian component with variable amplitude and FWHM.

Since different physical processes can give rise to emission lines with different FWHM, we perform two different simulations:

- **SIMUL A**, lines with FWHM=5 \AA , mainly due to star formation. The analysis of a spectrum of a typical star-forming galaxy from SDSS has shown a typical FWHM of the emission lines of $\sim 5 \text{ \AA}$.

- **SIMUL B**, lines with FWHM=9 \AA . Since, in the case of emission lines due to different processes than star formation, such as AGN activity, the lines are broader than the ones found for star-forming galaxies.

By changing the FWHM and the EW line in theoretical emission lines, we study the variation of the absorption line profile with the contamination of this different emissions. The results of the simulations (for the case of a SSP SED with age=10 Gyr) are shown in Fig 11. In both simulations (A and B) it is possible to identify a threshold of emission below which it is not possible to detect the H β emission component on the absorption line. We can appreciate the distortion of the H β line due to the introduction of an emission line, with increasing EW, only after $EW \leq -0.2 \text{ \AA}$ or $EW \leq -0.9 \text{ \AA}$ for SIMUL A or B, respectively. Emissions smaller than these values are not able to change the shape of the line, but however, they cause a intensity decrease in the absorption line, producing a drastic variation in the H β lick index measure, with $\Delta(H\beta)_{\text{max}} \sim 0.2 \text{ \AA}$ or $\Delta(H\beta)_{\text{max}} \sim 0.9 \text{ \AA}$ in SIMUL A and B, respectively. From this analysis, we can see that, for galaxies spectra with a resolution of the order of 3 \AA FWHM, there is a threshold detection limit of the H β emission line. As seen in previous section, the presence of [NII] λ 6584 emission in our samples may indicate the pres-

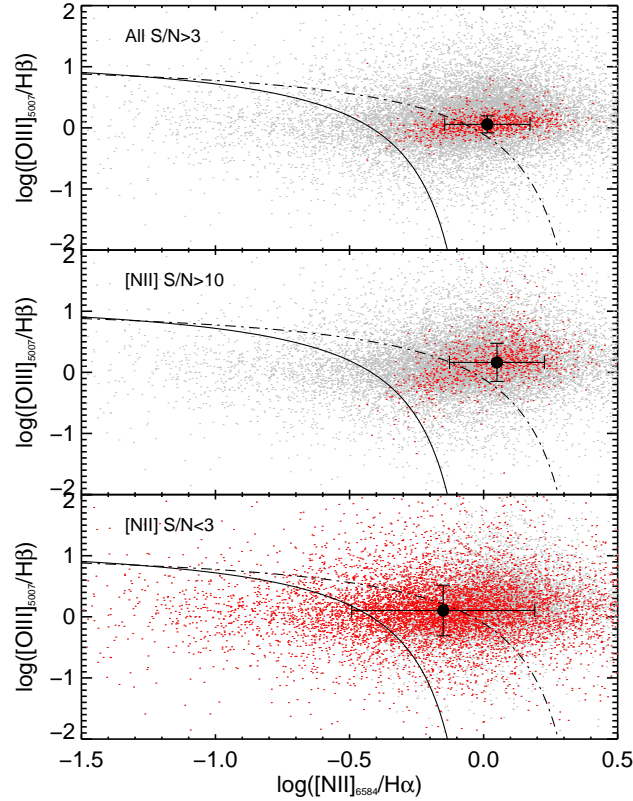


Figure 10. Distribution of the galaxies in the BPT diagram for the sample with all EW $S/N > 3$, EW([NII] λ 6584) $S/N > 10$ and EW([NII] λ 6584) $S/N < 3$, respectively from the top downwards. The gray cloud is the total sample. The dashed curve is the theoretical demarcation of Kewley et al. (2001), that separates star-forming galaxies and composites from AGN. The solid curve indicate the empirical division between pure star-forming galaxies from composite and AGN of Kauffmann et al. (2003). The median value and dispersion of the distribution are shown with the black symbols, in each subsample. (A color version of this figure is available in the online journal.)

ence of a weak $H\beta$ emission line with $EW \geq -0.28 \text{ \AA}$ for the sample with $S/N([NII]\lambda 6584) < 3$ and $EW \geq -0.9 \text{ \AA}$ for the sample with $S/N([NII]\lambda 6584) > 10$. These values are near to the threshold limit identified by simulations, therefore consistent with not being detectable in emission in our spectra. However, this potential hidden emission line can explain the detected offset between observational data and theoretical models discussed in the Section 3.2.

Performing the same simulations for SSP models with different ages (5 and 13 Gyr), we find that the detection limit for the $H\beta$ emission line does not change significantly.

4.5 Ionization sources

We find that even the less contaminated passive galaxies (sample with EW([NII] λ 6584) $S/N < 3$) may present an emission line contamination in the $H\beta$ absorption line.

As discussed in Section 4.3 using the BPT diagrams, we found that by selecting samples with higher S/N (sample with all lines $S/N > 3$ and sample with EW([NII] λ 6584) $S/N > 10$), our galaxies show typically LINERs-like rather than star-formation emission line ratios. These results are consistent with several imaging and spectroscopic observations (Phillips et al. 1986; Goudfrooij 1999; Sarzi et al. 2006; Annibali et al. 2010; Yan & Blanton 2012), which show a strong indication of the presence of warm gas (typically with $T \sim 10^4 \text{ K}$) in ETGs, with line ratios typically classified as LINERs. Despite the number of studies, two fundamental questions remain open: what is the origin of this ISM in the ETGs and what is the physical mechanism that determines its ionization?

Studies carried out on the kinematics of the gas, in most cases, show a misalignment between gas and stellar compo-

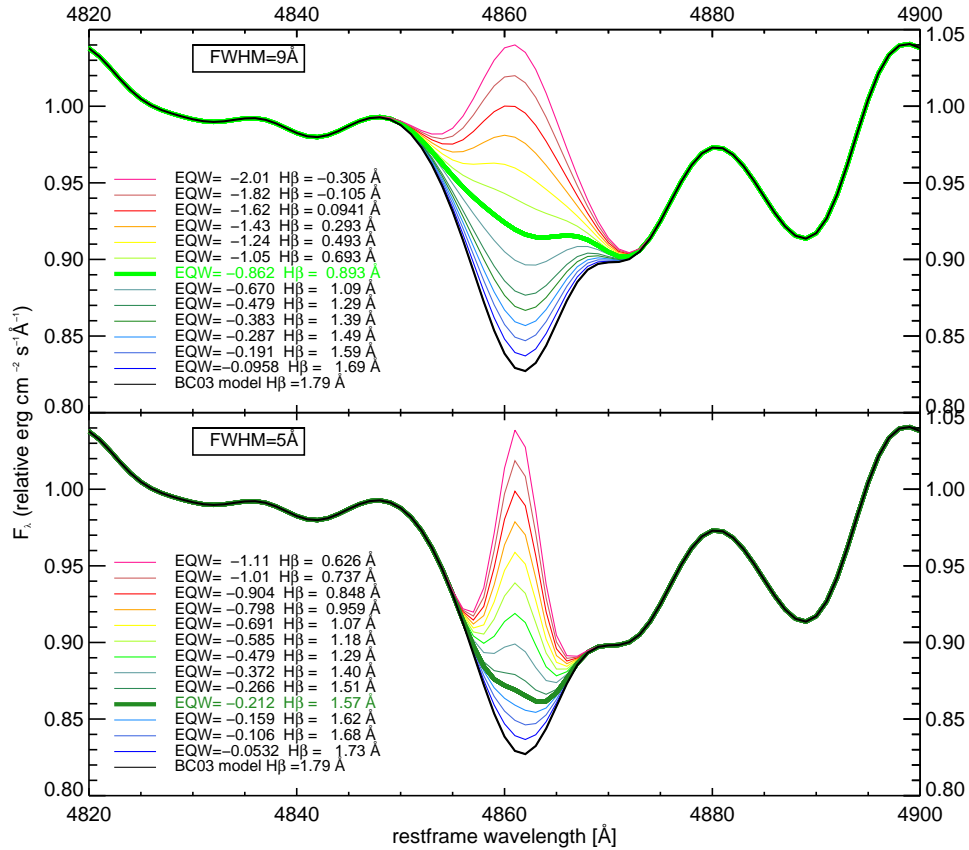


Figure 11. H β emission line simulations. We show the emission line effect within the H β absorption line (Simul A in the lower panel and Simul B in the upper panel). The different EW contaminations are plotted according with the labels (with different colors in the online journal). In both figures we can identify a threshold limit above which it is possible to detect an emission contamination (thick curves). The absorption model is a SSP with $Z/Z_\odot = 1$, age = 10 Gyrs and $\sigma = 200 \text{ km s}^{-1}$.

ment, suggesting an external origin of the gas (Caon et al. 2000). This observational evidence is in agreement with recent estimates of the gas emission metallicity, which seems to indicate an external source (Annibali et al. 2010).

However, as reported in Sarzi et al. (2006), the angular momenta measured do not seem to be consistent with a purely external origin.

The nature of the mechanism of ionization in LINERs is still not clear, but currently, the most accredited are photo-ionization by AGN activity (e.g. Ho 1999, 2000; Kewley et al. 2006), photo-ionization by post-asymptotic giant branch (post-AGB) stars (Binette et al. 1994) and fast shocks (e.g. Dopita & Sutherland 1995). The AGN activity, in fact, produces a source of energetic photons (X-rays and UV) able to ionize the interstellar medium; this is confirmed by some observed bubbles in LINERs galaxies (Baldi et al. 2009). For some objects, the presence of

AGN activity was also confirmed by observations of the radio core and X-rays point sources in their centers and UV variability (e.g. Ho 2008). Nevertheless, recent works have shown that in many cases the gas emissions are not only neighboring to the center, but an extended emission component also exists (Sarzi et al. 2006; Annibali et al. 2010; Yan & Blanton 2012; Singh et al. 2013). For this reason, the AGNs are not considered as the main cause of the LINERs-like emission lines in ETGs in favor of spatially extended ionizing sources, that in some cases follow the profile of stellar density (Yan & Blanton 2012). These properties are found in another source proposed for the first time by Binette et al. (1994): the post-AGB stars in the old stellar populations. These stars, after the AGB phase, are often surrounded by material ejected during their thermal pulses. After that, this material is dispersed in the ISM, and the AGB stars, having very high temperatures ($\sim 10^5 \text{ K}$), are ca-

pable of producing a diffuse radiation field. Recent spatially resolved spectroscopy observations obtained with a prototype of MANGA instrument, confirm the presence of an extended LINERs-like emission associated with spectral features of old and metal rich stars, see Belfiore et al. (2015). This scenario is also supported by the fact that post-AGBs are able to reproduce the variation of the parameter of ionization (defined as the ratio of the ionizing photon flux density to the electron density) with the radius of the galaxy (Yan & Blanton 2012). As expected in Binette et al. (1994) and Cid Fernandes et al. (2009), the post-AGBs are able to produce weak emission lines: $H\alpha \sim -0.6, -1.7 \text{ \AA}$ for a stellar population of 8 and 13 Gyrs respectively. This is very interesting because if this emission is expressed in terms of $H\beta$ emissions, similar to what was seen in the Section 4.3, we find a measure of $EW(H\beta) \sim -0.2, -0.6 \text{ \AA}$ emission line perfectly consistent with what is expected for our sample, especially with regard to the subsample with $EW([NII]\lambda 6584) S/N < 3 \text{ \AA}$. However, as shown in Yan & Blanton (2012), the post-AGB stars, while representing the ideal candidate for the creation of an photo-ionizing source, according to the most recent models, produce an ionization parameter too small compared to what is required by the observations. It should be noted that the current knowledge of the number density of post-AGB stars is still very uncertain, and therefore the question if these stars may be the main cause of emissions in older populations remains open. Finally, the emission lines with LINERs-like ratios can be produced by fast shock. However, this mechanism can not be mainly responsible for the emission in ETGs, since, as suggested in Yan & Blanton (2012), it provides emission lines with similar velocities between the different elements, contrary to what emerges from the recent estimates of velocity in different lines.

The data analyzed here do not give a direct indication of how to disentangle the three possible photo-ionization source candidates. However the observed anti-correlation between the $[NII]$ emission and the mass (see Sec. 4.1, Fig. 8) could disfavor the AGN hypothesis as a photo-ionization source, since its presence should instead be correlated with the galaxy mass (see Best et al. 2005; Brusa et al. 2009).

Finally, it is particularly important that the detection of emission lines in our sample seems to suggest a more frequent presence of ionized gas in these galaxies. This observation has significant effects on knowledge of this particular class of galaxies. The ETGs are not the simplest systems, but they are more complex in which only a knowledge of both the stellar contribution and of the ISM (and of the link between the two components) can shed light on their formation history and evolution.

5 CONCLUSIONS

In this paper, we explore the properties of the $H\beta$ Lick index of massive and passive ETGs, to estimate its robustness as an age indicator, being the Lick index more sensitive to the stellar population age and less affected by stellar metallicity, as suggested since the pioneering work of Worthey et al. (1994). The aim of this work is to establish its reliability as “cosmic chronometer” to trace the age evolution of the Universe as a function of redshift and to provide new constraints on the age of formation and evolution of galaxies stellar populations, eventually allowing measurement of the Hubble parameter $H(z)$ through the “cosmic chronometers” approach (Jimenez & Loeb 2002; Moresco et al. 2012a; Moresco 2015; Moresco et al. 2016a).

Using photometric and spectroscopic information, we select the most massive, passive and red ETGs in the SDSS-DR6 survey. The final sample consists of about 23914 galaxies with stellar mass $\log(M/M_\odot) > 10.75$ in the redshift range $0.05 \leq z \leq 0.3$. We divide the sample into four mass subsamples ($\Delta \log(M/M_\odot) = 0.25$) in order to avoid possible biases due to mass downsizing effect, and obtained four homogeneous subsamples in redshift of formation. All the samples are further divided into redshift bins and the spectra analyzed to measure the $H\beta$ Lick index, obtaining a median $H\beta - z$ relation in each mass subsample.

The main results of this analysis may be summarised as follows.

- Despite the rather strict selection criteria, all the spectra present characteristic features of a passive population, we find in the median stacked spectra of each mass bin clear ev-

idence of a weak [NII] emission line, with equivalent width of the order of -0.6 \AA , which can be interpreted as a hint of the presence of ionized gas. This residual emission line contamination may have a significant impact on the $H\beta$ Lick index measures. To address this issue, we split our samples on the base of the [NII] line, having a purer sample of 13626 galaxies selected with $\text{SN}([\text{NII}]) < 3$, and a more contaminated sample with $\text{SN}([\text{NII}]) > 10$. The analysis of these two subsamples confirmed the first hypothesis of a possible hidden contamination, with the sample with $\text{SN}([\text{NII}]) > 10$ presenting clear emission lines also in $H\alpha$, [OIII] and $H\beta$. We also find that the amplitude of the [NII] emission line anti-correlates with stellar mass.

- The analysis of all mass subsamples reveals a clear evolution of the $H\beta$ Lick index as a function of redshift, with the index decreasing with cosmic time. These trends are qualitatively consistent with a passive evolution, and are proven to be independent on many effects that can affect the analysis. These trends are also confirmed on both the original samples, and on the “purest” one obtained after selecting only galaxies with $\text{SN}([\text{NII}]) < 3$.

- At each redshift more massive galaxies present a median $H\beta$ index lower than less massive ones, confirming a mass-downsizing scenario for which more massive systems have assembled their stars earlier and faster. This result has been demonstrated not to depend on a selection effect due to the different velocity dispersions and metallicities of the samples; also in this case, this trend is found in both the original and the “purest” samples.

- The comparison with SPS models highlights an inconsistency with observable data, for which observed galaxies appear, in most cases, to be older than the age of the Universe at the given redshifts. These differences are greater than $\Delta H\beta \sim 0.1 \text{ \AA}$ and seems not reconcilable with any possible further re-selection of the samples. Only a stellar metallicity systematically higher than the one found in these samples ($Z/Z_{\odot} \sim 2$) may alleviate the tension between the data and models.

- We tested the presence of a weak $H\beta$ emission line in our stacked spectra by using GANDALF code and the emission corrections for the hydrogen features proposed by

Serven & Worthey (2010). We find that the recovered emission lines are very uncertain, detected with a very low confidence level and are model dependent. The median emission equivalent width ranges from ~ -0.1 to $\sim -0.27 \text{ \AA}$.

- We also find that in the stacked spectrum obtained from the “purest” sample with $\text{SN}([\text{NII}]) < 3$, there is a residual [NII] emission line contamination, even if very weak ($\text{EW} \sim -0.26$).

- Throughout simulations, we demonstrate that it exists a threshold limit below which an emission line component within the $H\beta$ absorption features would not be detectable; this threshold depends on the FWHM of the line, being $\text{EW}(H\beta) \leq -0.9 \text{ \AA}$ or $\leq -0.2 \text{ \AA}$ respectively for a broader and narrower FWHM (characteristic of star formation or AGN activity). We note that this EW is compatible with the observed offset between models and data.

In order to obtain a quantitative estimate of the age of formation of the galaxies and to test the feasibility of using this index as a “cosmic chronometers” indicator, the observed $H\beta$ -z relations should be calibrated on SPS models. However, the dependence of the normalization and of the slope of these relations on the different possible selections, and the apparent inconsistency with theoretical models, makes $H\beta$ an index that is difficult to rely on to estimate both absolute and relative ages, and this, as discussed, is due to a possible contamination of the line by an undetectable emission component. We discuss the possible candidates of ionization sources, finding a better agreement with post-AGB and slightly higher tension with AGN.

Despite this index has been historically identified as the best suited to constrain the age of a galaxy population, all the highlighted issues do not allow to use it for an accurate estimate of the absolute and differential age evolution of the stellar populations with redshift. Therefore, it is not possible to place constraints on the galaxies redshift of formation, and even the use of the $H\beta$ -z relation in the “cosmic chronometers” approach appears inadequate.

Finally, higher SN and resolution spectra may help in mitigating these problems during the selection phase, and to better disentangle a narrow emission line component for a

less biased measurement. Another possible option is to study higher order Balmer lines ($H\gamma$, $H\delta$), as e.g. suggested by the work of Vazdekis & Arimoto (1999), since those lines should be less affected by an underlying emission component. This analysis will be further exploited in a following paper.

ACKNOWLEDGEMENTS

The authors would like to thank the referee, Guy Worthey, for useful suggestions. We also thank Marco Mignoli, Claudia Maraston, Daniel Thomas, Jonas Johansson, Marcella Brusa and Gianni Zamorani for the stimulating discussions. We acknowledge Anna Gallazzi and Jarle Brinchmann for the SDSS data availability and their clarifications. AC is also grateful to Annalisa Citro and Salvatore Quai for sharing a preview of their results on SDSS stacked spectra analysis. Part of this work was supported by the INAF, Osservatorio Astronomico di Bologna and by the Dipartimento di Fisica e Astronomia, Università degli Studi di Bologna. We acknowledge the grants ASI n. I/023/12/0 "Attività relative alla fase B2/C per la missione Euclid" and PRIN MIUR 2010-2011 "The dark Universe and the cosmic evolution of baryons: from current surveys to Euclid". ACi, LP and MM acknowledge the PRIN MIUR 2015 "Cosmology and Fundamental Physics: illuminating the Dark Universe with Euclid".

REFERENCES

- Annibali F., Bressan A., Rampazzo R., et al., 2010, *A&A*, 519, A40
- Baldi A., Forman W., Jones C., et al., 2009, *ApJ*, 707, 1034
- Beifiori A., Maraston C., Thomas D., Johansson J., 2011, *A&A*, 531, A109
- Belfiore F., Maiolino R., Bundy K., Thomas D., 2015, *MNRAS*, 449, 867
- Best P. N., Kauffmann G., Heckman T. M., et al., 2005, *MNRAS* 362, 25-40
- Binette L., Magris C. G., Stasińska G., et al., 1994, *A&A*, 292, 13
- Borch, A., Meisenheimer, K., Bell, E. F., et al., 2006, *A&A*, 453, 869
- Bruzual G., Charlot S., 2003, *MNRAS*, 344, 1000
- Baldwin J. A., Phillips M. M., Terlevich R., 1981, *PASP*, 93, 5
- Brusa M., Fiore F., Santini P., et al. 2009, *A&A*, 507, 1277
- Bundy K., Ellis R. S., Conselice C. J., 2005, *ApJ*, 625, 621
- Bundy K., Ellis R. S., Conselice C. J., 2006, *ApJ*, 651, 120
- Burstein D., Faber S. M. et al, 1984, *ApJ*, 287, 586
- Caon N., Macchetto D., Pastoriza M., 2000, *ApJS*, 127, 39
- Cappellari M., Emsellem E. , 2004, *PASP*, 116, 138
- Choi J., Conroy C., Moustakes J., et al. 2014, *ApJ*, 792, 95
- Cid Fernandes R., Mateus A. et al., 2005, *MNRAS*, 358, 363
- Cid Fernandes R., Schlickmann M., Stasinska G., et al., 2009, *ASPC*, 408, 122
- Cimatti A., Daddi E., Renzini A., 2006, *A&A*, 453, L29
- Cimatti A., Cassata, P., Pozzetti, L. et al. 2008, *A&A*, 482, 21
- Citro A., Pozzetti, L., Moresco M., Cimatti A., 2015, *A&A*, 592, A19
- Cowie L. L., Songaila A., Barger A. J., 1999, *AJ*, 118, 603
- Dopita M. A., Sutherland R. S., 1995, *ApJ*, 455, 468
- Dunlop J., Peacock J., Spinrad H., et al., 1996, *Nature*, 381, 581
- Gallazzi A., Charlot S., Brinchmann J., et al., 2005, *MNRAS*, 362, 41
- Goudfrooij P., 1999, *ASPC*, 163, 55
- Graves G. J., Faber S. M., Schiavon R. P. and Yan R., 2007, *ApJ*, 671, 243
- Hanuschik R. W., 2003, *A&A*, 407, 1157
- Heavens A., Panter B., Jimenez R., et al., 2004, *Nature*, 428, 625
- Heckman T. M., 1980, *A&A*, 87, 152
- Ho L. C., 1999, *ApJ*, 510, 631
- Ho L. C., Rudnick G., Rix H.-W., et al., 2000, *ApJ*, 541, 120
- Ho L. C., 2008, *ARA&A*, 46, 475
- Hoaglin D. C., Mosteller F., Tukey J. W., 1983, John Wiley & Sons, pg. 404-414
- Ilbert O., Arnouts S., McCracken H. J., et al., 2006, *A&A*, 457, 841
- Ilbert O., Salvato M., Le Floc'h E., et al., 2010, *ApJ*, 709, 644
- Jimenez R., Loeb A., 2002, *ApJ*, 573, 37
- Jimenez R., Verde L., Treu T., et al., 2003, *ApJ*, 593, 622
- Kauffmann, G., Heckman T. M., White S., et al., 2003, *MNRAS*, 341, 33
- Kewley L. J., Dopita M. A., Sutherland R. S., et al., 2001, *ApJ*, 556, 121
- Kewley L. J., Groves B., Kauffmann G., et al., 2006, *MNRAS*, 372, 961
- Kochanek C. S., Eisenstein D. J., et al., 2012, *ApJS*, 200, 8
- Komatsu E., Smith K. M., Dunkley J., et al., 2011, *ApJS*, 192, 18
- Lee H.-c., Worthey G., 2005, *ApJS*, 160, 176

- Maraston C., Strömbäck G., 2011, MNRAS, 418, 2785
- Maraston C., Pforr J., Henriques B. M., Thomas D., 2013, MNRAS, 435, 2764
- Moresco M., Jimenez R., Cimatti A., et al., 2011, JCAP, 3, 45
- Moresco M., Cimatti A., Jimenez R., et al., 2012, JCAP, 8, 6
- Moresco M., Verde L., Pozzetti L., et al., 2012, JCAP, 7, 53
- Moresco M., Pozzetti L., Cimatti A., et al., 2013, A&A, 558, A61
- Moresco M., 2015, MNRAS, 450, L16
- Moresco M., Pozzetti L., Cimatti A., Jimenez R., et al. 2016, JCAP,05,014
- Moresco M., Jimenez R., Verde L. et al. 2016, JCAP, 12, 39
- Morton D. C., 1991, ApJS, 77, 119
- Osterbrock D. E., 1989, NYASA, 571, 99
- Pozzetti L., Bolzonella M., Zucca E., et al., 2010, A&A, 523, A13
- Riess A. G., Macri L., Casertano S., et al., 2011, ApJ, 730, 119
- Sarzi M., Falcón-Barroso J., Davies R. L., et al., 2006, MNRAS, 366, 1151
- Serven J., Worthey G., 2010, AJ, 140, 152
- Simon J., Verde L., Jimenez R., 2005, PhRvD, 71, 123001
- Singh R., et al. 2013, A&A, 558, A43
- Spinrad H., Dey A., Stern D., et al., 1997, ApJ, 484, 581
- Stern D., Jimenez R., Verde L., et al., 2010, JCAP, 2, 8
- Thomas D., Maraston C., Bender R., et al., 2005, ApJ, 621, 673
- Thomas D., Maraston C., Schawinski K., et al., 2010, MNRAS, 404, 1775
- Thomas D., Maraston C., Johansson J., 2011, MNRAS, 412, 2183
- Trager S. C., Worthey G., Faber S. M., et al., 1998, yCat, 211, 60001
- Tremonti C., Heckman T., Kauffmann G., et al., 2004, ApJ, 613, 898
- Worthey G., Faber S. M., Gonzalez J. J., et al., 1994, ApJS, 94, 687
- Yan R., Blanton M. R., 2012, ApJ, 747, 61
- Vazdekis A., Arimoto N., 1999, ApJ, 525, 144
- Zhang C., Zhang H., Yuan S., et al., 2012, arXiv:1207.4541
- Zucca E., Ilbert O., Bardelli S., et al., 2006, A&A, 455, 879

This paper has been typeset from a \LaTeX file prepared by the author.